

Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults

Catherine L. Sebastian,^{1,2} Nathalie M. G. Fontaine,³ Geoffrey Bird,^{2,4} Sarah-Jayne Blakemore,² Stephane A. De Brito,¹ Eamon J. P. McCrory,¹ and Essi Viding^{1,2}

¹Clinical, Educational and Health Psychology, Division of Psychology and Language Sciences, University College London, 26 Bedford Way, London, WC1H 0AP, UK, ²Institute of Cognitive Neuroscience, Division of Psychology and Language Sciences, University College London, 17 Queen Square, London, WC1N 3AR, UK, ³Department of Criminal Justice, Indiana University, 302 Sycamore Hall, Bloomington, Indiana, 47405, USA, and ⁴Department of Psychological Sciences, Birkbeck College, University of London, Malet Place, London, WC1E 7HX, UK

Theory of Mind (ToM) is the ability to attribute thoughts, intentions and beliefs to others. This involves component processes, including cognitive perspective taking (cognitive ToM) and understanding emotions (affective ToM). This study assessed the distinction and overlap of neural processes involved in these respective components, and also investigated their development between adolescence and adulthood. While data suggest that ToM develops between adolescence and adulthood, these populations have not been compared on cognitive and affective ToM domains. Using fMRI with 15 adolescent (aged 11–16 years) and 15 adult (aged 24–40 years) males, we assessed neural responses during cartoon vignettes requiring cognitive ToM, affective ToM or physical causality comprehension (control). An additional aim was to explore relationships between fMRI data and self-reported empathy. Both cognitive and affective ToM conditions were associated with neural responses in the classic ToM network across both groups, although only affective ToM recruited medial/ventromedial PFC (mPFC/vmPFC). Adolescents additionally activated vmPFC more than did adults during affective ToM. The specificity of the mPFC/vmPFC response during affective ToM supports evidence from lesion studies suggesting that vmPFC may integrate affective information during ToM. Furthermore, the differential neural response in vmPFC between adult and adolescent groups indicates developmental changes in affective ToM processing.

Keywords: Theory of Mind; empathy; adolescence; development; fMRI

INTRODUCTION

Theory of Mind (ToM) refers to the ability to attribute mental states to self and others, including knowledge, beliefs and intentions (Premack and Woodruff, 1978). Other researchers have also included understanding feelings within the definition of ToM (Shamay-Tsoory *et al.*, 2005). It is likely that ToM is a multidimensional process, requiring the integration of a number of components (Amodio and Frith, 2006). One recent model (Shamay-Tsoory *et al.*, 2010; Figure 1) distinguishes cognitive from affective sub-processes of ToM. Cognitive ToM refers to the ability to make inferences about beliefs and motivations, while affective ToM refers to the ability to infer what a person is feeling. According to this model, cognitive ToM is a prerequisite for affective ToM, which also requires intact empathy processing

(an ability to share and understand the emotional states of others; Singer *et al.*, 2009). Successful affective ToM processing, therefore, requires the integration of cognitive ToM and empathy. The first aim of the current study was to investigate the neural bases of cognitive and affective ToM using fMRI. Given recent evidence of continued development of the neural bases of ToM between adolescence and adulthood (Blakemore, 2008), a second aim was to explore the nature of this development in more detail by including an adolescent comparison group and an affective ToM condition.

fMRI studies exploring the neural bases of ToM have identified a network of regions that are commonly co-activated when participants are asked to think about their own or others' mental states, including posterior superior temporal sulcus at the temporoparietal junction (pSTS/TPJ), temporal poles, precuneus and medial prefrontal cortex (mPFC) (Frith, 2007). However, these studies have not explored the distinction between cognitive and affective ToM. A series of lesion studies by Shamay-Tsoory and colleagues (Shamay-Tsoory *et al.*, 2005, 2006; Shamay-Tsoory and Aharon-Peretz, 2007) has shown that patients with lesions to ventromedial PFC (vmPFC) show impairment on ToM tasks involving an affective component (e.g. understanding the affective state behind an ironic remark), but

Received 22 October 2010; Accepted 8 March 2011

Advance Access publication 4 April 2011

We are grateful to Dr Birgit Völlm for sharing her original stimuli, to Dr Simone Shamay-Tsoory for permitting us to present her original model in Figure 1, to Dr Darrick Jolliffe for his permission to use the Basic Empathy Scale, to the Birkbeck-UCL Centre for Neuroimaging for its support and to Andrea Greenwood (artist), Dr Alice Jones, John Rogers, Rosie Oakley and Isabel White for assistance at various stages of the project. This work was supported by the British Academy (53229), the UCL Department of Clinical, Educational and Health Psychology, and the Social Sciences and Humanities Research Council of Canada (post-doctoral fellowship to N.M.G.F.).

Correspondence should be addressed to Catherine Sebastian, Clinical, Educational and Health Psychology, University College London, 26 Bedford Way, London, WC1H 0AP, UK. E-mail: c.sebastian@ucl.ac.uk

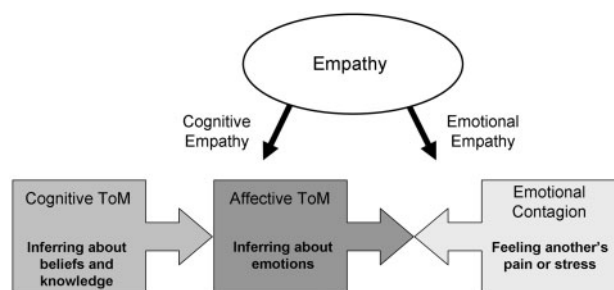


Fig. 1 The model of the relationship between cognitive ToM, affective ToM and empathy proposed by Shamay-Tsoory *et al.* (2010). Cognitive ToM is a prerequisite for affective ToM, which also requires cognitive and emotional (or affective) aspects of empathy. Figure adapted with author's permission.

do not show impairment on similar ToM tasks without an affective component (e.g. understanding the motivation behind an ironic remark in a non-affective context). Due to the fact that the lesions studied were not anatomically discrete, the reported damage in the vmPFC group extended well into both the medial and orbital PFC. Nevertheless, Shamay-Tsoory and Aharon-Peretz (2007) note that the region of greatest damage in this sample was the ventral portion of the medial PFC. These findings support the view that vmPFC may be required for affective but not cognitive ToM. This is a plausible hypothesis given that this region is well placed to integrate cognitive and affective information due to its extensive connections with regions involved in affective processing including the amygdala, temporal pole and anterior insula (Shamay-Tsoory *et al.*, 2006). Basic emotional (or affective) processing such as emotion perception and recognition is not an explicit component of the Shamay-Tsoory model, but we would argue that it is necessary (although not sufficient) for empathy processes which contribute to successful affective ToM (Figure 1).

Evidence from fMRI regarding the role of the vmPFC in affective ToM is currently mixed. Using drawings depicting emotional *vs* neutral social scenarios, Krämer *et al.* (2009) found activation in vmPFC [Tal: $-3\ 38\ -9$]. Simply viewing social scenarios (two people) relative to a single person activated the typical ToM network (including a more dorsal region of mPFC), but did not activate vmPFC. However, this study did not explicitly require participants to mentalize; nor did it directly compare cognitive and affective ToM. In another study Völlm and colleagues compared the neural response during cartoons requiring cognitive ToM (understanding intentions) *vs* cartoons requiring empathy (Völlm *et al.*, 2006). A conjunction analysis showed that both tasks activated the ToM network. The empathy condition activated mPFC (BA 10) to a greater extent than did the cognitive ToM condition; however, the peak of this activation fell in the dorsal portion of BA 10 [Tal: $9\ 53\ 14$], and as such cannot truly be considered to be vmPFC.

The current study aimed to test several predictions arising from the model of cognitive and affective ToM put forward by Shamay-Tsoory *et al.* (2010). The use of fMRI with

healthy participants allowed several predictions to be tested that are inaccessible to lesion methods. For example, the model predicts that cognitive and affective ToM should activate similar structures, but that affective ToM should additionally recruit regions hypothesized to integrate cognitive and affective information (including empathy processing) in order to predict outcomes (such as mPFC and vmPFC; Amodio and Frith, 2006; Shamay-Tsoory *et al.*, 2007), as well as those subserving empathic responding [such as insula (Singer *et al.*, 2009) and amygdala (Völlm *et al.*, 2006)]. Testing these predictions in a sample of healthy participants was the first aim of the present study.

The second aim was to explore the development of cognitive and affective ToM between adolescence and adulthood. There is some behavioural evidence that the development of cognitive ToM precedes that of affective ToM. For example, while children can pass second-order false belief tasks (understanding what person A understands about what person B thinks) from the age of 6 or 7 years (Perner and Wimmer, 1985), the ability to represent what person A understands about what person B feels (for example, in the understanding of social faux pas) appears later, between the ages of 9 and 11 years (Baron-Cohen *et al.*, 1999). Developmental fMRI studies further suggest that the neural substrates of ToM continue to develop during adolescence, long after children are able to perform complex cognitive and affective ToM tasks. For example, a meta-analysis by Blakemore (2008) found that adolescents activated mPFC (BA 10) to a greater extent than did adults on a number of tasks requiring inferences about mental states: both cognitive ToM, e.g. understanding intentions (Blakemore *et al.*, 2007), and affective ToM, e.g. irony comprehension (Wang, *et al.*, 2006) and understanding social emotions (Burnett *et al.*, 2009). However, no previous fMRI study has explicitly looked at similarities and differences in the neural processing of cognitive and affective ToM between adolescence and adulthood. Therefore, it is unclear whether the previously observed differential activation of mPFC between adolescence and adulthood is attributable to late development of a process that is shared between cognitive and affective ToM, or whether it reflects particularly protracted development in regions subserving more complex ToM demands, requiring the integration of cognitive ToM with empathy processing.

The third aim of the present study was to explore the relationship between ToM and self-reported ability to empathize, since the model by Shamay-Tsoory *et al.* (2010) suggests that affective ToM requires the integration of cognitive ToM and empathy. However, empathy is itself not a unitary construct, and there is considerable disagreement as to how the different dimensions of empathy should be defined. For the purposes of the present study, we explore potential relationships between cognitive and affective ToM, and cognitive and affective (or emotional) dimensions of empathy. We use the definitions of cognitive/affective

empathy adopted by Jolliffe and Farrington's (2006) Basic Empathy Scale (BES), with cognitive empathy defined as 'understanding another's emotions' and affective empathy as 'affect congruence', i.e. sharing another's emotions and being aware that the other person is the source of one's own affective state (de Vignemont and Singer, 2006). In adopting these definitions, the BES attempts to avoid confusing cognitive empathy with perspective taking (taking another's point of view), and affective empathy with sympathy (concern for the target person). Since this scale has not been used previously in relation to ToM ability, we aimed to characterize potential relationships, particularly between empathy subscales and affective ToM.

The current study used a cartoon vignette paradigm with separate conditions for cognitive ToM, affective ToM and a physical causality control condition. This allowed neural responses to each ToM subtype (cognitive and affective) to be explored relative to a control condition that did not require mental state attribution (physical causality), as well as to each other. The cartoon paradigm, adapted from Völlm *et al.* (2006), was chosen because it makes few demands on reading ability. As discussed above, Völlm and colleagues originally investigated the distinction and overlap between neural responses to ToM *vs* empathy. However, the cartoon scenarios used in the present study more closely followed the cognitive/affective ToM dichotomy as outlined by Shamay-Tsoory *et al.* (2010). Cognitive ToM cartoons required participants to understand beliefs or intentions, while affective ToM cartoons required participants to understand feelings, in order to select the appropriate story ending.

We tested only males for several reasons. First, although no previous studies have investigated the neural processing of cognitive *vs* affective ToM in an adolescent group, several studies investigating ToM processing in adolescents exist, but some have focused on females only (e.g. Blakemore *et al.*, 2007; Burnett *et al.*, 2009). Therefore, more data on the development of this ability in males is needed. Second, differing trajectories of structural brain development between males and females during adolescence (Giedd *et al.*, 1999; Raznahan *et al.*, 2010) suggests that averaging results across both sexes might result in noisy data that are not representative of either sex. Finally, prominent conditions associated with difficulties in either cognitive or affective ToM (e.g. autism spectrum disorders and psychopathy) are more commonly reported in males (Blair *et al.*, 2005; Baird *et al.*, 2006). Data on typical neural circuitry underpinning these abilities in healthy males, therefore, provides a backdrop for potential future studies on disordered populations.

We predicted that, across all participants, cognitive and affective ToM cartoons would elicit responses in the 'classic' ToM network (pSTS/TPJ, precuneus, temporal poles, mPFC), but that affective ToM would be associated with a greater or additional response in regions integrating cognitive and affective information (mPFC and vmPFC) and

those involved in basic affective processing (e.g. amygdala) and empathy (e.g. insula). We further predicted that developmental differences between adolescents and adults in the neural processing of ToM would be most pronounced for affective ToM, in line with previous behavioural evidence suggesting that affective ToM may develop later, and consistent with the Shamay-Tsoory model suggesting that affective ToM requires the integration of cognitive ToM and empathy. Finally, we conducted exploratory analyses to assess whether the neural processes associated with cognitive and affective ToM would be differentially related to cognitive and affective components of empathy as measured by self-report.

METHODS

Participants

Originally, 16 adolescents and 15 adults were recruited. One adolescent was subsequently excluded from all analyses due to excessive motion during MRI scanning. The final sample therefore comprised 15 participants in each group. Mean age was 14.18 years (*s.d.* = 1.88, range = 11.17–16.30) for the adolescent group and 28.88 years (*s.d.* = 4.54, range = 24.14–40.71) for the adult group. There were no group differences in relation to full scale IQ as measured by the Wechsler Abbreviated Scale of Intelligence (WASI; 1999) two-subtest version [adolescents: *M* = 106.40, *s.d.* = 10.60; adults: *M* = 111.73, *s.d.* = 7.70; *t*(28) = 1.58, *P* = 0.13]. No participants had a history of neurodevelopmental or psychiatric disorder, based on parental report (adolescents) or self-report (adults). Procedures were approved by the University College London Research Ethics Committee.

Experimental task

The task was adapted from a cartoon vignette paradigm used by Völlm *et al.* (2006). The scenarios were rewritten and the cartoons redrawn so that the scenarios would be easily understood by both adolescents and adults. All conditions were matched in terms of the number of protagonists per cartoon, and the complexity of the scenes. Stimuli consisted of 30 cartoons; each cartoon consisted of three frames telling a story, and one final screen with two choices of ending. The participant was asked to decide the appropriate ending.

The cartoons were divided into three conditions of 10 cartoons each; Affective ToM, Cognitive ToM and Physical Causality (PC). Each cartoon scenario portrayed two people to control for social content. For Affective ToM cartoons, participants were required to infer how one story character would react to their companion's affective state, in order to choose the correct ending. Cognitive ToM cartoons required an inference to be made based on the intentions or beliefs of one story character and their companion. PC scenarios required an understanding of cause and effect (e.g. sunshine melting snow or a football breaking a window): selecting the correct answer did not

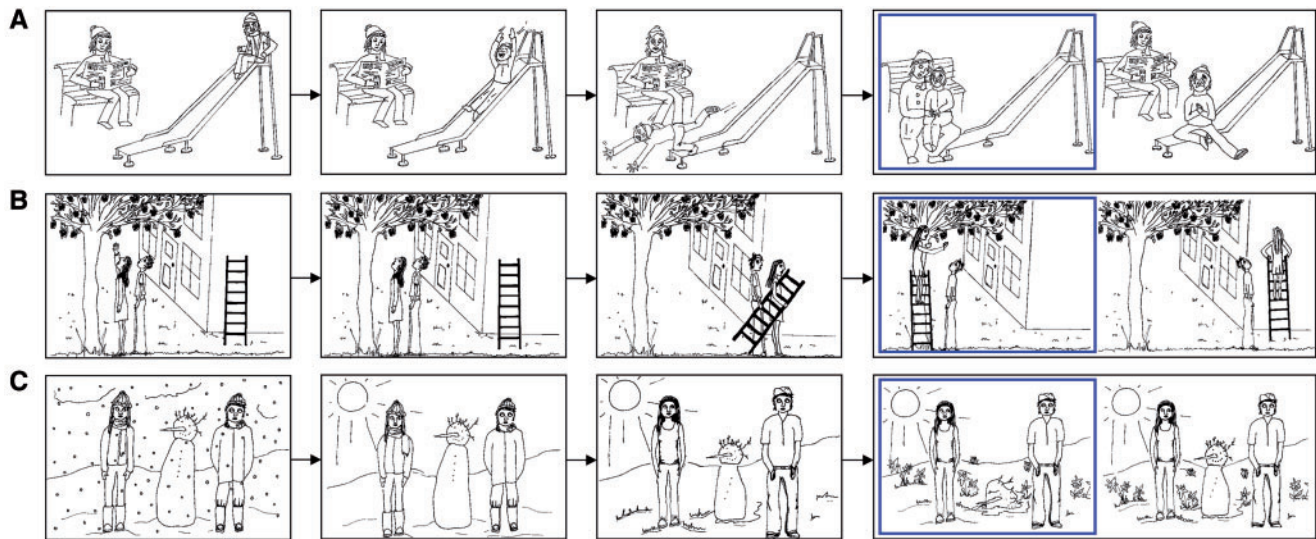


Fig. 2 Examples of the cartoon story stimuli for (A) Affective ToM, (B) Cognitive ToM and (C) Physical Causality conditions. Each frame of the story was sequentially displayed for 2 s. The choice between two endings was displayed for 5 s. A blue frame highlighted the participant's choice from the onset of the key press response until the end of the 5 s display. For illustrative purposes the correct answer is shown highlighted on the left of the display, although during the task the location of the correct answer was randomized.

require the understanding of mental states of the story characters appearing in the cartoon (Figure 2).

The 30 cartoons were presented in sets of 6 cartoons. In between sets, a fixation cross was displayed for 15 s. Each set consisted of two cartoons from each condition. These two cartoons were yoked together, although the order of the three conditions within a set was randomized. The order in which cartoons were presented was randomized anew for each participant. Each trial (i.e. presentation of one cartoon) started with an instruction screen displayed for 3 s: 'What happens next?' This was followed by the three 'story' frames, each presented for 2 s (6 s in total). The choice of endings was then displayed for 5 s. During this interval, the participant selected their chosen ending using a key press response. Their choice was highlighted with a blue box around the edge of the cartoon, and this remained onscreen until the end of the 5 s decision period. There was then an ISI of 1 s. Each trial therefore lasted a total of 15 s, and each block of two cartoons from each condition lasted 30 s. The task was piloted on children of the target age group (10–16 years) for both clarity of the stories and the appropriateness of the timings.

Questionnaire measures

Affective and cognitive empathy

The Basic Empathy Scale (BES; Jolliffe and Farrington, 2006) was used to assess affective (11 items) and cognitive (9 items) empathy by self-report. Example items included 'I get caught up in other people's feelings easily' (affective empathy) and 'I can often understand how people are feeling even before they tell me' (cognitive empathy). Each item was rated on a 5-point scale ranging from 'strongly disagree' (1)

to 'strongly agree' (5). The BES has demonstrated good validity (Jolliffe and Farrington, 2006). For the adolescent participants, the internal consistency (Cronbach's α) of the affective and cognitive empathy scales was 0.80 and 0.76, respectively. For the adult participants, the internal consistency of the affective and cognitive empathy scales was 0.86 and 0.90, respectively.

fMRI data acquisition

A Siemens Avanto 1.5 T MRI scanner was used to acquire a 5.5 min 3D T1-weighted structural scan, and multislice T2*-weighted echo planar volumes with BOLD contrast. The T2* EPI sequence used the following acquisition parameters: 46 2mm slices acquired in a descending trajectory with a 1 mm gap, TE = 50 ms; TR = 4100 ms; slice tilt = -30° ($T > C$); flip angle = 90° ; field of view = 192 mm; matrix size = 64×64 . Functional data were acquired in two runs of 9 min, with 134 volumes acquired per run.

fMRI data analysis

Imaging data were analysed using SPM8 (www.fil.ion.ucl.ac.uk/spm). The first four functional image volumes from each run were discarded to allow for T1 equilibrium effects, leaving 130 image volumes per participant. Pre-processing included rigid-body transformation (realignment) and normalization into the standard space defined by the Montreal Neurological Institute (MNI) template with a voxel size of $2 \times 2 \times 2$ mm, and smoothing with a Gaussian filter of 8-mm full width at half-maximum.

Across both runs of the cartoon task, a block analysis was conducted in order to compare neural activity associated with Affective ToM, Cognitive ToM and PC. For each run,

the time series of 130 image volumes was deconstructed into five block types, each of which was included as a separate regressor in the design matrix. These consisted of Affective ToM, Cognitive ToM and PC (variables of interest), as well as visual fixation and a variable comprising the times during which the instruction screen and ISI were displayed (variables of no interest). These five regressors were modelled as boxcar functions convolved with a canonical haemodynamic response function. The six realignment parameters were modelled as effects of no interest, in order to account for any variance due to head movement. For three adolescent participants, extra regressors were included to model a small number of corrupted images resulting from excessive motion. These images (one image for two participants and two images for one participant) were removed and the adjacent images interpolated in order to prevent distortion of the between-subjects mask. Data were high-pass filtered at 128 s to remove low-frequency drifts.

At the first level, four contrasts of interest were conducted for each participant: (i) Affective ToM > PC, (ii) Cognitive ToM > PC, (iii) Affective ToM > Cognitive ToM and (iv) Cognitive ToM > Affective ToM. The resulting contrast images were then entered into separate second-level analyses for each contrast of interest, where Group (adults, adolescents) served as a between-subjects variable in independent samples *t*-tests. Main effects of each of the four basic contrasts, as well the interaction between Condition and Group for each contrast, could then be explored. For the Condition by Group interaction analysis, the SPM was initially thresholded at $P < 0.001$, uncorrected. Reported regions are those reaching cluster-level significance at $P < 0.05$, FWE corrected, or those with $k \geq 5$ voxels which survived small volume correction at $P < 0.05$, FWE corrected with a 10 mm sphere centred on peak co-ordinates taken from the main effects analysis (an orthogonal contrast). Correlations were also conducted between BOLD signal associated with the four contrasts of interest and (i) BES subscales (BES affective and BES cognitive) as well as (ii) age in the adolescent group. The same thresholds were applied for these analyses as for the Condition by Group interactions.

RESULTS

Behavioural data

Task performance

For each participant, mean reaction times for correct trials (RTs) and percentage error rates were averaged across the two experimental runs. These data are displayed in Table 1. A Condition (Affective ToM, Cognitive ToM, PC) by Group (adults, adolescents) ANOVA with mean RT as the dependent variable showed no main effects of Condition or Group (P 's > 0.40), and no significant interaction ($P = 0.083$). In a Condition by Group ANOVA with percentage error rate as the dependent variable, there were no main effects of Condition or Group (P 's > 0.24). However, there was a

Table 1 Means and standard deviations for RT (ms) and percentage error data for the cartoon task, presented by Condition and Group

	Adults	Adolescents
Mean RT (s.d.)		
Affective ToM	1979 (253)	1967 (483)
Cognitive ToM	1963 (306)	1966 (461)
Physical causality	2006 (299)	1836 (428)
Percent errors (s.d.)		
Affective ToM	5.33 (6.40)	13.00 (11.77)
Cognitive ToM	5.67 (6.23)	8.33 (7.48)
Physical causality	8.00 (10.14)	4.33 (5.30)

significant interaction between Condition and Group: $F(2, 56) = 4.92$, $P = 0.011$. Post hoc tests showed that this was driven by a greater error rate in the Affective ToM condition in the adolescent group compared with both the adult error rate for Affective ToM (mean difference = 7.67%, $P = 0.035$) and the adolescent error rate for PC (mean difference = 8.67%, $P = 0.018$, Bonferroni corrected). Missed trials were excluded from these analyses. Missed trial rates were very low (below 1.5% across all Conditions and Groups), and an ANOVA showed no main effects or interactions involving Condition or Group (P 's > 0.60).

BES affective and BES cognitive subscales

For adolescents, the mean (s.d.) of the affective and cognitive empathy scales was 30.40 (6.30) and 35.47 (4.81), respectively. For adults, the mean (s.d.) of the affective and cognitive empathy scales was 37.15 (5.79) and 36.13 (4.47), respectively. Mean scores on the cognitive empathy scale did not differ between adolescent and adult groups [$t(28) = -0.39$, $P = 0.70$]. However, adolescents had lower scores of affective empathy compared with adults [$t(28) = -3.06$, $P = 0.005$].

fMRI data

Main effects of Condition

Regions reaching cluster-level significance at $P < 0.05$, FWE corrected, for the four contrasts of interest are shown in Table 2. For Affective ToM > PC, significant clusters were found in bilateral mPFC/vmPFC, precuneus/posterior cingulate cortex, pSTS/TPJ and temporal poles (Figure 3A). For Cognitive ToM > PC, significant clusters were found in bilateral precuneus/posterior cingulate cortex and temporal poles, and the right pSTS/TPJ (Figure 3B). For Affective ToM > Cognitive ToM, significant clusters were found in bilateral precuneus/posterior cingulate cortex and right temporal pole. The predicted result in vmPFC for this contrast did not survive cluster-level correction; however, a cluster of 30 voxels in right vmPFC reached significance at $P < 0.001$, uncorrected at the voxel level (peak voxel [4 52 -8], $t = 4.15$, $z = 3.63$).

Table 2 Regions showing a main effect at $P < 0.05$ with FWE correction at the cluster level for contrasts Affective ToM > PC, Cognitive ToM > PC, Affective ToM > Cognitive ToM and Cognitive ToM > Affective ToM

Brain region	BA	L/R	Peak voxel	<i>k</i>	<i>z</i>	Cluster corrected, <i>P</i> -value
Affective ToM > PC						
Precuneus/posterior cingulate cortex	31, 7	R	8 -54 26	2382	6.81	<0.001
	23	L	-6 -46 26		5.40	
STS/TPJ ext. occipitotemporal cortex	40, 39, 19	L	-54 -50 18	883	6.24	<0.001
Medial/ventromedial prefrontal cortex	9, 10	R	10 50 22	948	5.66	<0.001
	10	L	-8 54 0		4.29	
Temporal pole	38, 21	R	50 8 -24	1326	5.62	<0.001
STS/TPJ ext. occipitotemporal cortex	40, 19	R	52 -48 20	1404	5.06	<0.001
Temporal pole	21	L	-62 -10 -14	426	4.73	<0.001
Cognitive ToM > PC						
Precuneus/posterior cingulate cortex	31, 7	R	8 -60 26	7555	7.73	<0.001
ext. occipitotemporal cortex	19	L	-40 -80 22		6.02	
STS/TPJ	40, 39	R	50 -48 20	1337	4.82	<0.001
Temporal pole	21, 22	R	56 -6 -16	480	4.69	<0.001
Temporal pole	21	L	-62 -12 -12	138	4.45	0.043
Affective ToM > Cognitive ToM						
Precuneus/posterior cingulate cortex	23, 7	L	-6 -52 26	280	4.34	0.002
	31	R	4 -54 28		3.89	
Temporal pole	38, 21	R	50 10 -24	195	3.97	0.015
Cognitive ToM > Affective ToM						
Inferior temporal gyrus	37	L	-46 -54 -6	17 121	7.30	<0.001
ext. parahippocampal gyrus	19	R	30 -44 -10		7.04	
ext. precuneus	7	L	-20 -76 50		6.86	
Premotor cortex	6	L	-22 8 52	1130	6.05	<0.001
Premotor cortex	6	R	24 8 52	1045	5.49	<0.001
Premotor cortex	6	L	-40 -8 30	466	4.42	<0.001
ext. Dorsolateral prefrontal cortex	9	L	-46 4 36		3.91	

BA = Brodmann area. Where more than one BA is shown, the peak voxel falls in the first BA, but the cluster extends to include the others listed. L/R = left/right, peak voxel = MNI *xyz* co-ordinates, *k* = cluster size.

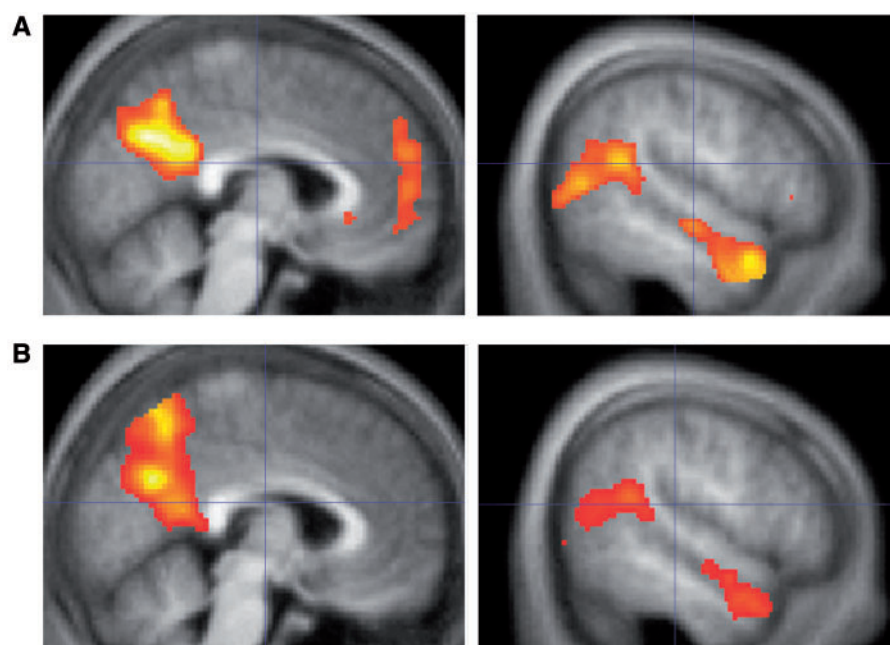


Fig. 3 Main effects relative to Physical Causality for (A) Affective ToM and (B) Cognitive ToM. Medial (left panel) and right lateral (right panel) views are shown at a threshold of $P < 0.001$, overlaid on an average structural scan from all 30 participants. For both contrasts, significant clusters were seen in the pSTS/TPJ, temporal poles and precuneus. Affective ToM > PC was additionally associated with BOLD response in bilateral mPFC/vmPFC.

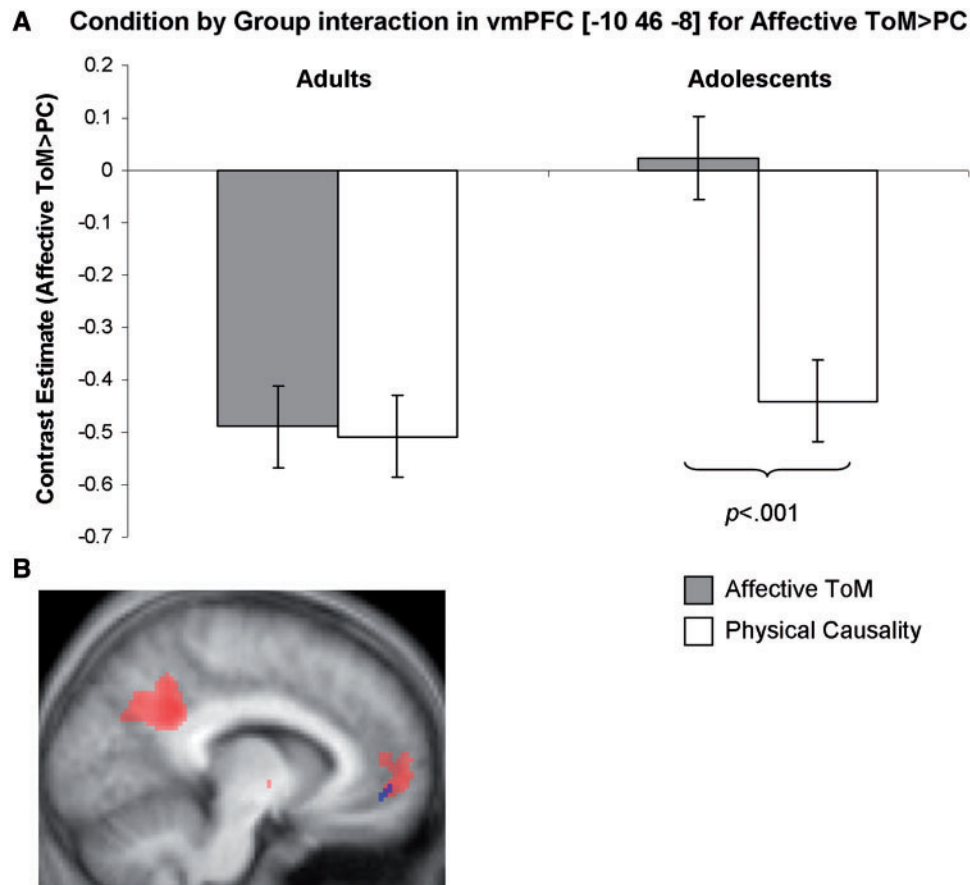


Fig. 4 Group difference in BOLD response in the vmPFC for the contrast Affective ToM>PC: **(A)** Graph showing the shape of the interaction in the peak voxel [-10 46 -8]. Adolescents showed a significantly greater response in this region during Affective ToM than during PC, while the adult group showed no difference between conditions. Please note that some responses are plotted as deactivations relative to fixation baseline. This does not change the interpretation of relative differences between conditions; **(B)** An overlay showing the main effect across all participants in left vmPFC for Affective ToM>PC (red blob) as well as the cluster showing an interaction with age group (blue blob). Age group differences are evident for the most ventral part of the vmPFC activation only. Data are displayed at a threshold of $P < 0.001$, overlaid on an average structural scan from all 30 participants.

The reverse contrast Cognitive ToM > Affective ToM was associated with significant clusters in bilateral inferior temporal gyrus and premotor cortex.

Interactions with group

For the Affective ToM > PC contrast, an interaction between Condition and Group in left vmPFC (adolescents > adults) survived SVC with a 10 mm sphere centred on the co-ordinate from the main effect peak in this region [-8 54 0] (peak voxel for the interaction = [-10 46 -8]; BA 32/10; $k = 7$; $z = 3.54$; SVC $P = 0.045$, FWE-corrected) (Figure 4). Post hoc t -tests conducted on the peak voxel showed that the interaction was driven by the adolescent group, with a greater response during the Affective ToM condition than during PC [$t(28) = 5.93$, $P < 0.001$]; in contrast, there was no difference between conditions in the adult group [$t(28) = 0.23$, $P = 0.41$; all one tailed]. As adolescents made more errors than did adults during the Affective ToM cartoons relative to the PC cartoons, this analysis was also conducted with the error rate difference included as a

covariate. Results were very similar to the original analysis: peak voxel = [-10 46 -8]; BA 32/10; $k = 6$; $z = 3.50$; SVC $P = 0.043$, FWE. There were no interactions with Group for any other contrasts of interest.

Correlations

Age

For correlations conducted within the adolescent group only, there was a significant positive relationship between age and BOLD response for the Cognitive ToM > PC contrast in the left occipitotemporal cortex (peak voxel = [-38 -84 26]; BA = 19; $k = 28$; $z = 3.78$; $P = 0.019$, SVC with a 10 mm sphere using the co-ordinate [-40 -80 22] from the Cognitive ToM > PC main effect). There were no further correlations with age for any of the other contrasts of interest.

BES-affective empathy scale

Correlations with this measure were conducted across the whole sample (Table 3). The group difference in

Table 3 Regions showing a correlation between the BES-Affective Empathy subscale and BOLD response for the contrasts of interest across all participants

Brain region	BA	L/R	Peak voxel	k	z	FWE-corrected <i>P</i> -value
BES-affective (Affective ToM > PC)						
Positive relationship—none						
Negative relationship						
Premotor cortex	6	R	26 −6 50	192	5.20	0.009
Ventrolateral prefrontal cortex	47	R	44 40 −8	316	4.96	<0.001
	46	R	42 32 16		4.42	
	46/47	R	44 36 4		3.91	
Cerebellum, posterior lobe	—	L	−32 −60 −40	389	4.93	<0.001
		L	−22 −66 −46		4.85	
		L	−32 −70 −30		3.80	
BES-affective (Cognitive ToM > PC)						
Positive relationship						
Parahippocampal gyrus ext. hippocampus	—	L	−22 −18 −24	182	4.95	0.012
	—	L	−12 −18 −24			
Negative relationship—none						
BES-affective (Affective ToM > Cognitive ToM)^a						
Positive relationship—none						
Negative relationship						
Post-central gyrus	40	L	−58 −28 20	234	4.83	0.004
Ventrolateral prefrontal cortex	45	R	56 14 2	248	4.21	0.003
ext. insula	13	R	46 12 2		4.05	
ext. superior temporal gyrus	22	R	62 0 4		3.79	
Posterior insula	13	L	−40 −8 4	593	4.11	<0.001
	13	L	−38 −2 10		4.08	
ext. superior temporal gyrus	22	L	−50 2 2		3.93	

^aCognitive ToM > Affective ToM correlations with BES are the reverse of those shown here. Results are significant at $P < 0.05$ FWE-corrected at the cluster level.

L/R = left/right, peak voxel = MNI xyz co-ordinates, k = cluster size.

self-reported BES-affective scores was controlled for by using independent samples *t*-tests across the two groups, and including the BES-affective subscale as a continuous independent variable which was modelled separately in each group. The effect of the BES-affective subscale, independent of the group difference in BES-affective score, could then be ascertained. For the Affective ToM > PC contrast, there was a negative relationship between BES-affective empathy score and BOLD response in the right ventrolateral PFC, right premotor cortex and left cerebellum. Similarly, for the Affective ToM > Cognitive ToM contrast, negative relationships were found in the right ventrolateral PFC, the left post-central gyrus and the left posterior insula. There were no regions showing a positive relationship with BES-affective score for either Affective ToM > PC or Affective ToM > Cognitive ToM. For Cognitive ToM > Physical Causality, there was a positive relationship between BES-affective Empathy score and BOLD signal in the left parahippocampal gyrus/hippocampus. No regions showed a negative relationship with this measure.

BES-cognitive empathy scale

There were no significant correlations involving this measure.

DISCUSSION

This study used a cartoon vignette paradigm to explore the distinction and overlap in the neural processing of cognitive and affective ToM. Relative to the Physical Causality condition, both Cognitive and Affective ToM elicited a response in regions classically associated with ToM, including bilateral pSTS/TPJ, precuneus and temporal poles. In support of the model by Shamay-Tsoory *et al.* (2010), the Affective ToM condition additionally recruited bilateral mPFC, extending both dorsally and ventrally, relative to the PC condition. Although it did not survive FWE correction, vmPFC responded to a greater extent during Affective ToM than during Cognitive ToM cartoons. Increased response in vmPFC in the adolescent group relative to adults during the Affective ToM condition (relative to PC) suggests that the role of vmPFC in mediating affective ToM continues to change between adolescence and adulthood. Finally, exploratory analyses across groups relating self-reported empathy to neural responses during the task revealed several regions in which the neural response to affective ToM was negatively correlated with self-reported affective empathy.

Behavioural data

Behavioural data showed that adolescents made significantly more errors on the Affective ToM cartoons than did adults.

However, there were no group differences for Cognitive ToM or PC cartoons. This finding suggests that Affective ToM is particularly challenging, and may continue to develop between adolescence and adulthood. It also fits with evidence showing that children need to be older to solve tasks requiring affective ToM (age 9–11 years for understanding social faux pas: Baron-Cohen *et al.*, 1999) than to solve tasks requiring cognitive ToM (age 6–7 years for second order cognitive ToM: Perner and Wimmer, 1985). The model by Shamay-Tsoory *et al.* (2010) suggests that affective ToM requires the integration of cognitive ToM and empathy (Figure 1). Recent studies have shown continued behavioural and neural development during adolescence in both cognitive ToM (Blakemore, 2008; Dumontheil *et al.*, 2009) and empathy (Decety and Michalska, 2010). Future work could address whether the relative immaturity of affective ToM in adolescence is due to immaturity in cognitive ToM and/or empathy components independently, or whether it results from difficulties integrating these two components. An additional factor to consider is that sub-components within the construct of empathy also develop at different rates (Decety, 2010). Of relevance to this issue, the current study found significantly lower self-reported affective empathy in the adolescent group relative to adults, but no group difference in self-reported cognitive empathy. This is somewhat surprising, since it has been suggested that the development of affective empathy precedes that of cognitive empathy (Decety, 2010), and this result will need replication in a larger sample.

Neural correlates of cognitive and affective ToM

fMRI responses to Cognitive and Affective ToM conditions relative to PC were broadly similar to each other, and included bilateral pSTS/TPJ, temporal pole and precuneus. These regions are classically associated with ToM (see Frith, 2007 and Saxe, 2006 for discussions of each region's proposed role in ToM processing). Their activation in both ToM conditions supports the proposal by Shamay-Tsoory *et al.* (2010) that an ability to infer mental states is required for both cognitive and affective ToM. Importantly however, only Affective ToM (relative to PC) activated a significant cluster in mPFC, extending both dorsally and ventrally. Ventromedial PFC (vmPFC) also responded in the contrast Affective ToM > Cognitive ToM, although this cluster did not reach significance at the FWE-corrected level. These findings provide some support for the idea that affective ToM requires additional processing compared with cognitive ToM, possibly reflecting the integration of cognitive and affective information in vmPFC (Shamay-Tsoory *et al.*, 2007, 2010).

However, the lack of a response during cognitive ToM in the more dorsal portion of mPFC could be considered surprising, given that many accounts of the role of mPFC in ToM are not affect specific. For example, it has been suggested that this region mediates meta-representational ability

(Amodio and Frith 2006; Frith and Frith, 2007), or that it decouples mental states from reality (Gallagher and Frith, 2003). However, several lesion studies have shown intact cognitive ToM when this region is damaged (Bird *et al.*, 2004; Shamay-Tsoory *et al.*, 2005, 2006; Shamay-Tsoory and Aharon-Peretz, 2007); thus it may be that mPFC is not necessary for all ToM tasks (Saxe, 2006). Another possibility is that matching all conditions for social content (two characters in each cartoon) eliminated an mPFC response in cognitive ToM relative to the PC condition, since previous studies have found that mPFC responds to the presence of two people relative to a single person even when the task does not require mentalizing (Krämer *et al.*, 2009).

One way in which the current findings deviate from the model by Shamay-Tsoory *et al.* (2010), is that the direct contrast between Affective and Cognitive ToM conditions in the present study suggested that both dimensions of ToM are associated with responses in distinct (as well as overlapping) networks. Affective ToM was associated with a greater response in bilateral precuneus and right temporal pole relative to Cognitive ToM (in addition to the uncorrected result in vmPFC), while the reverse contrast revealed activation in inferior temporal gyrus/occipitotemporal cortex, premotor cortex and dorsolateral PFC (dlPFC). Although not explicitly predicted, this latter result is of interest, since a recent repetitive transcranial magnetic stimulation (rTMS) study (Kalbe *et al.*, 2010) found that rTMS over the dlPFC selectively accelerated reaction times on a cognitive ToM task, but had no effect on affective ToM RTs. Thus, both the current findings and those of Kalbe *et al.* (2010) suggest that affective and cognitive ToM are each associated with their own distinct networks, in addition to regions that overlap between the two (pSTS/TPJ, temporal poles, precuneus). A similar conclusion was reached in a study by Völlm *et al.* (2006); and although the focus of that study was on comparing cognitive ToM and empathy, the definition of empathy was very similar to that of affective ToM used here and elsewhere.

Comparison of fMRI data between adolescents and adults

Comparison of fMRI data between adolescent and adult groups revealed a region of vmPFC that responded in adolescents but not in adults during the Affective ToM > PC contrast. This region overlaps with the region identified in lesion studies by Shamay-Tsoory and colleagues (2005, 2006, 2007) as mediating the integration of cognitive and affective information to enable successful affective ToM. fMRI data suggests that the functional neural correlates of ToM have a protracted developmental trajectory (Blakemore, 2008). However, the finding of a developmental difference during affective but not cognitive ToM in the present study may indicate that the trajectory of affective ToM is particularly protracted. The idea that affective ToM is more complex than cognitive ToM is further supported by the finding of

a higher error rate in the adolescent group than in the adults for the Affective ToM condition only. However, the fact that the result in vmPFC was still significant when error rates were included as a covariate suggests that error rate alone cannot explain the difference between adolescent and adult groups.

The nature of the interaction between age group and condition (Affective ToM > PC) suggests that the most ventral portion of the vmPFC cluster showing a main effect responded only in the adolescent group (Figure 3). At first glance this might seem somewhat paradoxical, i.e. adults activate a region that lesion evidence indicates is crucial for task performance to a lesser extent than do adolescents. However, it is possible that some complex cognitive processes become more automatic and rely less on prefrontal structures with increasing age. For example, there is evidence from studies of developmental lesion patients showing that vmPFC damage sustained early in life has a more detrimental effect on social and emotional functioning than similar lesions sustained in adulthood (Anderson *et al.*, 1999). Thus, structures such as vmPFC may be, particularly, important for the acquisition of social cognitive skills during development, but are less crucial to task performance in adulthood. The current study suggests that the process of acquiring complex social cognitive skills such as affective ToM continues into adolescence. However, the mechanisms by which neural responses are reduced and cognitive performance is enhanced between adolescence and adulthood is still unclear. It may be that age differences in cognitive strategy drive the observed difference in neural response. Alternatively, or in addition, physiological changes such as the pruning of excess synapses between adolescence and adulthood may increase efficiency of neural networks and drive cognitive development (see Blakemore, 2008 for a discussion of these issues).

Correlations with self-reported empathy

We also explored potential correlations between self-reported empathic abilities and BOLD response across age groups during the cartoon task. The pattern of results suggests that self-reported affective empathy correlates to a greater extent with BOLD response during affective ToM than does cognitive empathy. The lack of any correlations with the cognitive empathy subscale was somewhat surprising, since cognitive empathy and affective ToM are often defined in similar ways, i.e. in terms of 'understanding emotions' (Jolliffe and Farrington, 2006; Shamay-Tsoory *et al.*, 2010). This may be due to the relatively small variance obtained for the BES-cognitive subscale compared with the BES-affective subscale.

Significant correlations between the BES-affective empathy subscale and contrasts involving Affective ToM were negative, i.e. these brain regions responded to a greater extent for the contrasts involving Affective ToM in individuals reporting lower levels of affective empathy. In the case of frontal control regions, such as ventrolateral PFC,

individuals with greater self-reported empathy may have had less difficulty regulating an affective response to others' distress and misfortune, which may aid in taking an empathic stance. Alternatively, individuals low in affective empathy may automatically regulate emotion arising from another's distress to a greater extent than more empathic individuals: they may also have taken a more cognitive approach to the task, increasing the contribution of cognitive control regions such as ventrolateral PFC. It is less clear how the negative correlations in some of the other regions should be interpreted. As these analyses were exploratory, future studies should focus on designing a more direct test of the link between empathy and Affective ToM.

Some limitations should be mentioned in the context of the present study, which could be addressed by future work. First, the naturalistic nature of the cartoon stimuli meant that it was difficult to match the stimuli for visual complexity. Differences in visual complexity might explain the extensive activation in occipitotemporal cortex for the Cognitive ToM > Affective ToM contrast. Second, it is not yet clear how conceptual differences in complexity between affective and cognitive ToM, and differences in neural responses between these conditions, relate to differences in the complexity of neural computations performed. Third, this study focused on males only, and future work using this paradigm should also include females. However, there are also advantages in collecting data from males only; for example such data can provide an important starting point for exploring neural correlates of affective and cognitive ToM in disorders such as psychopathy and ASD, which are more prevalent in males.

An additional consideration is that, although predictions in the current study were based on the model of Shamay-Tsoory *et al.* (2010), there are alternative conceptualizations of the relationship between cognitive and affective ToM. For example, Brothers (1990) argues that affective cues (such as facial expressions) are used to infer motivations and intentions, while Frith and Frith (2007) suggest that affect sharing provides a mechanism for inferring another's goal or intention. Thus, both cognitive and affective ToM may rely on basic emotional processing, followed by reasoning about the inferred emotion or intention (Coricelli, 2005). This is in contrast to the model discussed in the current study, which suggests that making inferences about beliefs and knowledge (cognitive ToM), and affective processes such as empathy and emotional contagion, contribute separately to the ability to make inferences about emotions (affective ToM; Figure 1). It may be that non-affective mental state inferences are aided by basic affective processing. However, this does not alter the conclusions of the present study: namely that Affective ToM places additional demands relative to PC and Cognitive ToM, and this is particularly evident during adolescence as indexed by error rates and vmPFC response.

The current study explored the distinction and overlap in the neural processing of cognitive and affective ToM using a

non-verbal task with both adolescent and adult participants. To our knowledge, this is the first fMRI study to explicitly investigate the contrast between cognitive and affective ToM, hypothesized by Shamay-Tsoory and colleagues on the basis of evidence from lesion patients. Our findings suggest that both affective and cognitive ToM are associated with distinct neural substrates in addition to areas of common overlap in the traditional 'ToM network'. However, they also support the view that vmPFC plays a specialized role in affective as opposed to cognitive ToM. The current study also provides further evidence of functional development within the 'social brain' between adolescence and adulthood (Blakemore, 2008), with a greater BOLD response in adolescents than adults in a subregion of vmPFC during Affective ToM relative to Physical Causality. Thus, it appears that the neural basis of the ability to integrate affective information into ToM-based decisions continues to develop between adolescence and adulthood.

Conflict of Interest

None declared.

REFERENCES

- Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7(4), 268–77.
- Anderson, S.W., Bechara, A., Damasio, H., Tranel, D., Damasio, A.R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, 2(11), 1032–7.
- Baird, G., Simonoff, E., Pickles, A., et al. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, 368(9531), 210–5.
- Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, 29(5), 407–18.
- Bird, C.M., Castelli, F., Malik, O., Frith, U., Husain, M. (2004). The impact of extensive medial frontal lobe damage on 'Theory of Mind' and cognition. *Brain*, 127(4), 914–28.
- Blair, R.J.R., Mitchell, D.G.V., Blair, K.S. (2005). *The Psychopath: Emotion and the Brain*. Oxford: Blackwell.
- Blakemore, S.J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, 9(4), 267–77.
- Blakemore, S.J., Den Ouden, H., Choudhury, S., Frith, C. (2007). Adolescent development of the neural circuitry for thinking about intentions. *Social Cognitive and Affective Neuroscience*, 2(2), 130–9.
- Brothers, L. (1990). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, 1, 27–51.
- Burnett, S., Bird, G., Moll, J., Frith, C., Blakemore, S.J. (2009). Development during adolescence of the neural processing of social emotion. *Journal of Cognitive Neuroscience*, 21(9), 1735–50.
- Coricelli, G. (2005). Two-levels of mental states attribution: from automaticity to voluntariness. *Neuropsychologia*, 43, 294–300.
- Decety, J. (2010). The neurodevelopment of empathy in humans. *Developmental Neuroscience*, 32, 257–67.
- Decety, J., Michalska, K.J. (2010). Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. *Developmental Science*, 13(6), 886–99.
- de Vignemont, F., Singer, T. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences*, 10, 435–41.
- Dumontheil, I., Apperly, I.A., Blakemore, S.J. (2009). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, 13(2), 331–8.
- Frith, C.D. (2007). The social brain? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 671–8.
- Frith, C.D., Frith, U. (2007). Social cognition in humans. *Current Biology*, 17, R724–32.
- Gallagher, H.L., Frith, C.D. (2003). Functional imaging of 'Theory of Mind'. *Trends in Cognitive Sciences*, 7(2), 77–83.
- Giedd, J.N., Blumenthal, J., Jeffries, N.O., et al. (1999). Brain development during childhood and adolescence: a longitudinal MRI study [2]. *Nature Neuroscience*, 2(10), 861–3.
- Jolliffe, D., Farrington, D.P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, 29(4), 589–611.
- Kalbe, E., Schlegel, M., Sack, A.T., et al. (2010). Dissociating cognitive from affective theory of mind: a TMS study. *Cortex*, 46(6), 769–80.
- Krämer, U.M., Mohammadi, B., Doñamayor, N., Samii, A., Münte, T.F. (2009). Emotional and cognitive aspects of empathy and their relation to social cognition - an fMRI-study. *Brain Research*, 1311, 110–20.
- Perner, J., Wimmer, H. (1985). "John thinks that Mary thinks that..." attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, 39(3), 437–71.
- Premack, D.G., Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1, 515–26.
- Raznahan, A., Lee, Y., Stidd, R., et al. (2010). Longitudinally mapping the influence of sex and androgen signaling on the dynamics of human cortical maturation in adolescence. *Proceedings of the National Academy of Sciences*, 107(39), 16988–93.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, 16, 235–9.
- Shamay-Tsoory, S.G., Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia*, 45(13), 3054–67.
- Shamay-Tsoory, S.G., Harari, H., Aharon-Peretz, J., Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, 46(5), 668–77.
- Shamay-Tsoory, S.G., Tibi-Elhanany, Y., Aharon-Peretz, J. (2006). The ventromedial prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. *Social Neuroscience*, 1(3–4), 149–66.
- Shamay-Tsoory, S.G., Tomer, R., Berger, B.D., Goldsher, D., Aharon-Peretz, J. (2005). Impaired "affective theory of mind" is associated with right ventromedial prefrontal damage. *Cognitive and Behavioral Neurology*, 18(1), 55–67.
- Singer, T., Critchley, H.D., Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, 13, 334–40.
- Völlm, B.A., Taylor, A.N.W., Richardson, P., et al. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *NeuroImage*, 29(1), 90–8.
- Wang, A.T., Lee, S.S., Sigman, M., Dapretto, M. (2006). Developmental changes in the neural basis of interpreting communicative intent. *Social Cognitive and Affective Neuroscience*, 1(2), 107–21.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.